April 22, 2008

Dr. George Gray
Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

Dear Dr. Gray:

The Board of Scientific Counselors (BOSC) has completed a requested program review of the Office of Research and Development's (ORD) Science and Technology for Sustainability (STS) Research Program. A seven-member BOSC subcommittee, including one consultant to the subcommittee, was charged to conduct the program review. Dr. John Giesy, a member of the BOSC Executive Committee, chaired the STS Research Subcommittee that conducted the review. The BOSC Subcommittee members participated in five teleconferences and a 2-day face-to-face review meeting of the STS Program, which was held in April 2007. This review was a detailed retrospective and prospective analysis of the program and included extensive materials describing the STS Multi-Year Plan, individual research programs, and budgetary information. The BOSC Executive Committee reviewed the report and requested appropriate clarification and revision. The report was revised, vetted and approved for transmission to ORD by the Executive Committee in March 2008.

The principal objectives of the review were to evaluate: "(1) program relevance and quality; (2) program design and implementation; (3) progress achieved toward meeting long-term goals (LTGs); (4) stakeholder involvement and the degree to which research is consistent with needs articulated at regional and local levels; and (5) the degree to which research "outputs" are being used by stakeholders."

The LTGs of the STS Research Program seek to develop metrics and decision tools that are adopted to quantitatively assess and promote sustainability along with technologies to solve environmental problems contributing to sustainable outcomes. The summary findings of the review point to a Program that meets or exceeds expectations in achieving its LTGs relative to tools and technology development and their adoption. The creation and adoption of metrics for quantitative assessment of sustainability is in too early a stage for qualitative ranking at the time of review but every indication is

that the excellent research being conducted in the STS Research Program will contribute strongly to that goal.

As an outcome of the review, the BOSC Subcommittee developed recommendations for the ORD to consider based on the materials provided and the discussions organized as part of the review. The text of the report provides the full context and detail for these comments as well as other specific recommendations. This report is anticipated to further assist ORD in longer term program enhancement, comparative analysis with other programs, and intermediate research investment decision-making.

On behalf of the BOSC Executive Committee and the Science and Technology for Sustainability Research Subcommittee, it is my pleasure to transmit this program review report to ORD. The BOSC expects that the report will assist ORD in continuing to improve its science, and assist and inform clients within and outside the Environmental Protection Agency of the significance of the Agency's research and its utilization.

Please feel free to contact me if you have any questions concerning this report. We look forward to ORD's response.

Sincerely,

Gary S. Sayler
Chair

B·O·S·C

BOARD OF SCIENTIFIC COUNSELORS

REVIEW OF THE OFFICE OF RESEARCH AND DEVELOPMENT'S SCIENCE AND TECHNOLOGY FOR SUSTAINABILITY RESEARCH PROGRAM AT THE U.S. ENVIRONMENTAL PROTECTION AGENCY

Final Report

BOSC SUBCOMMITTEE ON SCIENCE AND TECHNOLOGY FOR SUSTAINABILITY RESEARCH

Dr. John P. Giesy (Chair), University of Saskatchewan
Dr. Wayne Landis, Western Washington University
Dr. Concepción Jiménez-González, CEHS GlaxoSmithKline
Dr. Earl R. Beaver, Institute for Sustainability
Dr. Martin Abraham, University of Toledo
Dr. Ted Tomasi, ENTRIX, Inc.

Subcommittee Consultant
Dr. Peter Blaze Corcoran, Florida Gulf Coast University

EPA Contact
Cloris Slocum, Designated Federal Officer

March 21, 2008

This report was written by the Science and Technology for Sustainability Research Subcommittee of the Board of Scientific Counselors, a public advisory committee chartered under the Federal Advisory Committee Act (FACA) that provides external advice, information, and recommendations to the Office of Research and Development (ORD). This report has not been reviewed for approval by the U.S. Environmental Protection Agency (EPA), and therefore, the report's contents and recommendations do not necessarily represent the views and policies of the EPA, or other agencies of the federal government. Further, the content of this report does not represent information approved or disseminated by EPA, and, consequently, it is not subject to EPA's Data Quality Guidelines. Mention of trade names or commercial products does not constitute a recommendation for use. Reports of the Board of Scientific Counselors are posted on the Internet at http://www.epa.gov/osp/bosc.

TABLE OF CONTENTS

I. SUMMARY

I.1 Background and Charge to the Subcommittee

The Executive Committee of the Board of Scientific Counselors (BOSC) of the Office of Research and Development (ORD) within the U.S. Environmental Protection Agency (EPA) has agreed to undertake a series of reviews of major ORD research programs. These evaluations conducted by the BOSC provide recommendations from independent experts that provide guidance to ORD to help:

- Plan, implement, and strengthen the program;
- Compare the program with programs designed to achieve similar outcomes in other parts of EPA and in other federal agencies;
- Make research investment decisions over the next 5 years;
- Prepare EPA's performance and accountability reports to Congress under the Government Performance and Results Act (GPRA); and
- Respond to assessments of federal research programs such as those conducted by the Office of Management and Budget (OMB) (OMB highlights the value of recommendations from independent expert panels in guidance to federal agencies.[1,2]).

Independent expert review is used extensively in industry, federal agencies, congressional committees, and academia. The National Academy of Sciences has recommended this approach for evaluating federal research programs.[3]

This report addresses ORD's Science and Technology for Sustainability (STS) Research Program. The BOSC Science and Technology for Sustainability Subcommittee conducted a prospective and retrospective review of the STS Research Program's relevance, quality, performance, and scientific leadership. The evaluation is based on the review of a large volume of written materials describing the STS Multi-Year Plan (MYP), individual ORD research programs, and budgetary information. The evaluation took place over the period January through December 2007 and included five conference calls and a 2-day site visit to the EPA facility in Cincinnati, Ohio, in April 2007. The site visit featured direct presentations by STS Research Program leaders and scientists and it provided ample time for the Subcommittee members to question EPA staff on program details.

The objectives of the review were to evaluate: (1) program relevance and quality; (2) program design and implementation; (3) progress achieved toward meeting long-term goals (LTGs);

(4) stakeholder involvement and the degree to which research is consistent with needs articulated at regional and local levels; and (5) the degree to which research "outputs" are being used by stakeholders. The research of STS is organized into three LTGs and this report provides an analysis of the STS Research Program in the context of each of these LTGs. The report then considers the STS Research Program in holistic terms and provides specific conclusions and recommendations.

I.2 Findings and Recommendations

Overall, the Subcommittee of the BOSC that conducted the review found that the STS Research Program is meeting expectations. Some elements of the Program are excellent and exceed expectations. Where the Program does not exceed expectations, the primary reason is that these program elements are small components and lack a critical mass of personnel. Other elements were difficult to judge because they currently are in transition to the future focus on sustainability.

One of the difficulties in assessing parts of the Program is that there is neither a standard definition of sustainability nor a clear framework for deriving an appropriate conceptual model. EPA has developed many frameworks in the past, notably the framework documents for risk assessment that were significant steps in the development of the field. Without a clear definition of sustainability or a means of deriving case-specific criteria, it will be very difficult to determine the efficacy of a specific metric or decision tool. A clear definition of sustainability also will be required to fund appropriate extramural research programs that will support the overall sustainability research program. An overall recommendation is that EPA make an effort to derive a clear definition of sustainability and a framework for its application to a broad range of human activities.

The STS Research Program has some excellent researchers who are world leaders in their fields and many others who are making solid contributions. The quality of the research is apparent and was accomplished with relatively limited resources. In particular, leveraging of available resources through partnerships with other agencies, both local and federal, allows the STS Research Program to achieve more than it otherwise could. The availability of resources, however, often directs the type of studies that are undertaken; whereas the research should be directed by the researchers' determinations of the critical questions in the field. The overall findings and recommendations are given by LTG. A more detailed discussion and interpretation of the findings and suggestions is provided in the body of the report.

Long-Term Goal 1 (LTG 1): Metrics—Decision-makers adopt ORD-identified and developed metrics to quantitatively assess environmental systems for sustainability.

Because LTG 1 is only now being developed and the Program reorganized, there was insufficient information to allow the Subcommittee to assign this LTG a qualitative score. The Subcommittee reviewed the plan, however, and provides some overall guidance below:

- ✍ Develop an outline for how metrics for sustainability will be developed. This should include criteria for assessing the utility and predictability of metrics.

- ✍ Coordinate metric development with other LTGs.

- ✍ Determine a strategy of how metrics will be used.

- ✍ Going forward, an extramural program based on the Technology for a Sustainable Environment (TSE) Program could be crafted to emphasize metrics and how technologies move toward improving the measures.

- ✍ Testing protocols should be established to determine if the metrics are measuring the intended functions, if they are consistent in their evaluation, if they are sufficiently independent, and if they can be effectively used to determine if specific actions are driving society to become more sustainable.

- ✍ Sustainability targets need to be identified so that appropriate metrics can be designed and tested.

- ✍ Critical experiments should be designed that allow testing of hypotheses within the realm of defined metrics.

- ✍ The predictability of the models should be evaluated and sensitivity analyses conducted.

- ✍ Evaluation of the metrics should be done systematically and quantitatively.

- ✍ A team that was better integrated throughout EPA could draw on additional resources that could enhance their effectiveness.

- ✍ There needs to be significant interaction between this LTG and, in particular, LTG 2, which are intimately tied together.

- ✍ LTG 1 metrics should be used to inform LTG 3 activities.

Long-Term Goal 2 (LTG 2). Decision Tools—Decision-makers adopt ORD-developed decision support tools and methodologies to promote environmental stewardship and sustainable environmental management practices.

Qualitative Score: ***Exceeds Expectations***

The Subcommittee determined that the STS Research Program was exceeding expectations relative to LTG 2. The Program is relatively mature in this area and a great deal of progress has been made. The progress toward achieving this LTG has been excellent and has had a large impact on the field of sustainability. The Subcommittee provides a few observations and suggestions below:

- The life cycle assessment (LCA) programs, metrics, and procedures developed under the Pollution Prevention and New Technologies (P2NT) Research Program are relevant and important to the goals of EPA, stakeholders, and the international community. The STS Research Program is positioned to move these initiatives forward and is encouraged to build on this strength.

- LTG 2 could be improved through targeted extramural collaborations on the development of new tools or cooperation on the advancement of existing tools or tools being developed in the private sector.

- Efforts should be made to reach a wider set of stakeholders, such as nongovernmental organizations (NGOs), state agencies, etc.

- The actual outputs and outcomes could be more clearly defined and communicated to targeted sectors.

Long-Term Goal 3 (LTG 3). Technologies—Decision-makers adopt innovative technologies developed or verified by ORD to solve environmental problems, contributing to sustainable outcomes.

Qualitative Score: ***Meets Expectations***

While the Subcommittee found the overall performance of the STS Research Program relative to LTG 3 to be meeting expectations, a range of performances was observed. Some Program elements were performing at a very high level and some would be classified as Exceeding Expectations; however, as a whole, the Subcommittee members thought that the overall performance was Meeting Expectations. Some observations and recommendations are given below:

- The People, Prosperity, and the Planet (P3); Small Business Innovation Research (SBIR); and Environmental Technology Verification (ETV) Programs all have been highly relevant to the mission of EPA and the elements in these programs should be preserved whenever possible.

- The relevance and impact of the Green Technology Program is less apparent and this program needs to be assessed internally to determine if it is serving a function that is not being met already by the private sector and academia.

- The P3 Program should improve the solicitation/judging criteria to require a clear statement by students as to the effects articulated via sustainability metrics or decision tools. A clear tie-in with the goals of LTG 1 and LTG 2 could be developed.

- More emphasis should be placed on measurement.

- The ETV Program should encourage an increased role in supporting emerging markets in trades/mitigation/offsets, such as mercury/greenhouse gases, etc.

- An analysis should be conducted to determine if there are emerging markets in this trade/offset area that have a barrier surrounding verification issues.

- The SBIR Program should increase its use of sustainability metrics in selection criteria and increase the linkage of program outcomes to sustainability metrics.

- Consideration should be given to redirecting the Green Technology Program or replacing it with an extramural grants program.

- Results derived from the Green Technology Program have not been effectively communicated to larger industrial enterprises.

- The Program could benefit from a more systematic evaluation of the program outcomes, such as tracking of careers of recipients to obtain information that can be used to measure impact as outcome.

II. INTRODUCTION TO THE PROCESS OF OBTAINING OVERALL FINDINGS AND RECOMMENDATIONS

II.1 Goals, Charge, and Structure of the Review

The Office of Research and Development (ORD) relies on expert external review to assess the scientific quality and performance of its research programs. In this report, the Science and Technology for Sustainability (STS) Subcommittee presents the findings of a review of ORD's STS Research Program. The STS Research Program is one of more than 10 research programs within ORD. ORD performs and funds research that provides technical standards and scientific information to support the U.S. Environmental Protection Agency's (EPA) broad mission of protecting the environment and, specifically, to support EPA's various regulatory functions.

Many "customers" and "stakeholders" utilize ORD's research products. Foremost among these are the regulatory divisions of EPA that are mandated to provide technical standards for regulatory decisions. These technical standards must be grounded in well-documented scientific knowledge. Other important customers include state regulatory entities, tribes, and local governments. The stakeholder population also includes nongovernmental organizations (NGOs) that have a stake in environmental conservation; the private sector that must adapt to and implement regulatory decisions that require clear and understandable technical standards; and finally, the general public that has a strong stake in EPA's mission of protecting human health and the environment. The combination of growing world population, rapidly expanding gross domestic product, and other forces will place significant stress on the Earth's resources and on humanity's ability to maintain or improve environmental quality. Today's generation and future generations thus face the challenge of preventing or mitigating the negative consequences that come with this population growth and economic expansion while simultaneously working to improve the protection of human health and environmental quality.

Because of the nature of research, it is not possible to measure the creation of new knowledge as it develops—or the pace at which research progresses or scientific breakthroughs occur. Demonstrating research contributions to outcomes is very challenging[4] when federal agencies conduct research to support regulatory decisions, and then rely on third parties[5]—such as state environmental agencies—to enforce the regulations and demonstrate environmental improvements. Typically, many years may be required for practical research applications to be developed and decades may be required for some research outcomes to be achieved.

In designing and managing its research programs, ORD emphasizes the importance of identifying priority research questions or topics to guide the research. Similarly, ORD recommends that its programs develop a small number of performance goals that serve as indicators of progress to answer the priority questions and to accomplish outcomes. Short-term outcomes are accomplished when research is applied by specific clients—to strengthen environmental decisions or regulations, for example. These decisions and resulting actions (e.g.,

the reduction of contaminant emissions or restoration of ecosystems) ultimately contribute to improved environmental quality and health.

In a comprehensive evaluation of science and research at EPA, the National Research Council[9] recommended that the Agency substantially increase its efforts to explain the significance of its research products and to assist clients inside and outside the Agency in applying them. In response to this recommendation, ORD has engaged science advisors from client organizations to serve as members of its research program teams. These teams help identify research contributions with significant decision-making value and help plan for their transfer and application.

For ORD's environmental research programs, periodic retrospective analysis at intervals of 4 or 5 years is needed to characterize research progress, to identify when clients are applying research to strengthen environmental decisions, and to evaluate client feedback about the research. Conducting program evaluation at this interval enables assessment of research progress, the scientific quality, and decision-making value of the research, and whether research progress has resulted in short-term outcomes for specific clients.

The STS Research Subcommittee of the Board of Scientific Counselors (BOSC) conducted this review (see Appendix A for a list of Subcommittee members). Prior to the review, the Subcommittee met twice via conference call (January 23, 2007 and February 21, 2007) for orientation to the federal review process and to discuss the review procedures. The review was conducted at the EPA facility in Cincinnati, Ohio, on April 25 and 26, 2007. Subsequent to the Cincinnati meeting, teleconferences were held on May 30, 2007, September 6, 2007, and December 11, 2007, to discuss the draft report. The review meeting in Cincinnati and three teleconferences were conducted as open meetings under the guidelines of the Federal Advisory Committee Act (FACA) (see Appendix B for a schedule of the Subcommittee meetings).

The format of this review is relatively new. The BOSC has recently begun an intensive review process for the various ORD activities and this is an evolving process. One impetus for the review is OMB's Program Assessment Rating Tool (PART) evaluation. The STS Research Program (formerly the Pollution Prevention and New Technologies [P2NT] Research Program) received a Results Not Demonstrated rating in its PART evaluation. Based on this evaluation, a restructuring of the Program and subsequent external review appeared in order to more fully assess the successes and failures of the restructured Program. A major purpose of this review is, therefore, to assist the STS Research Program in adapting to new requirements and expectations as it works to meet the needs of the nation. Specifically, in this review the BOSC implemented the assignment of non-numerical overall quality ratings. This rating tool had been used in mid-cycle reviews conducted by the BOSC, but this is the first program review where it was implemented. The ratings assigned in this review also reflected a recent change in the terminology applied by the BOSC such that the rating of Satisfactory was changed to Meets Expectations. The draft charge to the Subcommittee is provided in Appendix C.

A Brief Overview of the Science and Technology for Sustainability Research Program

As ORD's P2NT Research Program is nearing its completion, a new research program has been created, the STS Research Program, which is focused on the question of sustainability. Although this represents a new research direction for ORD, the STS Research Program will include a select group of research efforts that had their genesis within the P2NT Research Program. Key documents that lay the groundwork for this new Program are EPA's ORD Sustainability Research Strategy and the STS Multi-Year Plan (MYP), both of which were recently reviewed by EPA's Science Advisory Board (SAB). Research questions have been detailed in ORD's Sustainability Research Strategy, a document that first establishes the reasons why sustainability is an important issue, and then frames the Agency's approach by focusing on the issue of environmental sustainability. An extensive analysis of the relevant sustainability literature revealed six broad themes of environmental sustainability. They are:

1. Natural Resource Protection
2. Non-Renewable Resource Conservation
3. Long-Term Chemical and Biological Impacts
4. Human-Built Systems and Land Use
5. Economics and Human Behavior
6. Information and Decision-Making

The STS Research Program addresses a small subset of the environmental sustainability themes described above. The express purpose of the STS Research Program is to provide a foundation for subsequent sustainability efforts in the Agency. The STS Research Program, along with its collaborators and partners, will develop integrating decision support tools, sustainability metrics and indicators, and technologies that support sustainable outcomes, and in the process, provide technical support to broader regional and national sustainability policies and initiatives. This will be done through the pursuit of three long-term goals (LTGs):

- LTG 1 – Metrics: Decision-makers adopt ORD-identified and developed metrics to quantitatively assess environmental systems for sustainability.

- LTG 2 – Decision Tools: Decision-makers adopt ORD-developed decision support tools and methodologies to promote environmental stewardship and sustainable environmental management practices.

- LTG 3 – Technologies: Decision-makers adopt innovative technologies developed or verified by ORD to solve environmental problems, contributing to sustainable outcomes.

Several ongoing research efforts have been incorporated into the STS Research Program. These include:

- A Sustainable Environmental Systems Team that conducts multidisciplinary research will now be addressing long-term, systems-based solutions to the sustainable management of a regional ecosystem by using sustainability metrics and indicators.

- Life Cycle Assessment (LCA) research will continue to improve LCA methods through the development of streamlined approaches and the use of material flow methods.

- The ongoing Environmental Impact Modeling effort will expand the current suite of chemical impact models to include temporal- and spatially-based sustainability concerns.

- The in-house Green Chemistry Program, working through formal partnerships with industry, will emphasize the development of sustainable chemistries and technologies.

- The Environmental Technology Verification (ETV) Program, which provides quality controlled evaluation of new technologies, will incorporate sustainability metrics in all evaluations.

- The Small Business Innovation Research (SBIR) Program, which provides critical financial support to innovative technology developers, will emphasize the importance of sustainability in its solicitations.

- The annual People, Prosperity, and the Planet (P3) Student Sustainability Design Competition will continue its collegiate contest focused on promoting sustainable solutions to environmental concerns.

II.2 Program Assessment

The performance of each program was assessed for specific aspects of either the P2NT or STS Research Programs (as identified), or for all activities supporting the Program's LTGs. The Subcommittee developed answers to a series of questions provided by ORD relative to *Program Relevance, Program Structure, Program Quality, Scientific Leadership, Coordination and Communication,* and *Outcomes* and also provided a *Summary Assessment.*

The STS Research Program has three LTGs identified in the MYP. These three LTGs are:

- **LTG 1:** Identify and create scientifically-based sustainability metrics.

- **LTG 2:** Develop decision support tools that promote environmental stewardship and sustainable management practices.

- **LTG 3:** Develop, apply, and demonstrate innovative technologies that solve environmental problems and provide sustainable outcomes.

The assessment was based on questions given in the charge (see Appendix C) with respect to those P2NT Research Program elements that will be continued in the STS Research Program under LTGs 2 and 3. Because no metrics development work was conducted within the P2NT Research Program while the proposed plan was reviewed and suggestions made, no summary assessment was given for LTG 1.

One of the major difficulties in providing specifics in the following evaluations is the lack of a clear EPA-wide definition of sustainability or a supporting framework. Without these items, it is difficult to judge the exact relevance of a metric, analysis tool, or extramural program. A clear and precise definition of sustainability is critical to direct and evaluate the program.

Program Relevance

Factors Considered: To assess *Program Relevance*, the Subcommittee determined whether: (1) P2NT outputs were/are used by the Agency and stakeholders; (2) the STS Research Program addresses issues raised within the ORD Sustainability Research Strategy; (3) the STS MYP has clear goals and priorities; (4) stakeholders (e.g., program and regional offices) are involved in the planning and prioritization of the research; (5) outputs from the STS Research Program are likely to be used by stakeholders; and (6) the STS Research Program outlines a well-coordinated effort with outside research organizations, nationally and internationally, that will avoid duplication of effort and promote synergistic collaboration. Specifically, the Subcommittee answered the following questions:

- How relevant and consistent has P2NT research been with respect to Agency goals and customer needs?
- How evident are the public benefits of the P2NT research?
- How consistent are the LTGs of the STS Research Program with achieving the Agency's strategic plan?
- How responsive is the new STS Research Program direction to client needs and recommendations from outside advisory boards?

Program Structure

Factors Considered: To assess *Program Structure*, the Subcommittee determined whether: (1) the STS MYP addresses research questions identified in the Sustainability Research Strategy; (2) the MYP uses appropriate criteria to select research projects; and (3) the LTGs are appropriate for planning the research and for identifying long-term priorities that meet the scientific needs of the Agency and Program customers. Specifically, the Subcommittee answered the following questions:

- How clear and logical are the LTGs in the STS MYP for organizing and planning the research and demonstrating outcomes of the Program?
- How appropriate is the science used to achieve each LTG in the STS MYP? (i.e., Is the Program asking the right questions, or has it been eclipsed by advancements in the field?)
- To what extent does the STS MYP describe an appropriate flow of work that reasonably reflects the anticipated pace of scientific progress and timing of client needs?

- How logical is the STS Research Program design? How clearly identified are the STS Research Program priorities?

Program Quality

Factors Considered: To assess *Program Quality*, the Subcommittee considered whether: (1) the research involved a sufficient number of internal and external peer reviews and was modified in response to these reviews; (2) used state-of-the art science; and (3) used a competitive merit-based process that maintained quality when awarding extramural funds. Specifically, the Subcommittee answered the following questions:

- How good is the scientific quality of the P2NT research products?

- How appropriate is the science that has been used in the P2NT research?

- To what extent have appropriate means been employed to ensure quality P2NT research (including peer review, competitive funding, etc.)?

Scientific Leadership

Factors Considered: To assess the *Scientific Leadership*, the Subcommittee evaluated the degree to which ORD staff members are seen as leaders in the field and active participants in national and international science and technology professional bodies. Specifically, the Subcommittee reviewed and assessed the leadership role that ORD staff members have taken in contributing to advancing the current state-of-the-science for tools, methodologies, and technologies that support environmental decision-making.

Coordination and Communication

Factors Considered: To assess *Coordination and Communication*, the Subcommittee reviewed the extent to which the Program has collaborated with clients and stakeholders and leveraged resources from related research programs. Specifically, the Subcommittee answered the following questions:

- How effectively has ORD engaged outside organizations (both within and outside government) for the P2NT research?

- How effectively has ORD collaborated with and obtained input from others on research objectives, especially to avoid duplication of effort?

- How effective were the mechanisms used for communicating research results for the P2NT Research Program, both internally and externally?

Outcomes

Factors Considered: To assess *Outcomes*, the Subcommittee assessed the extent to which: (1) the Program's products met stakeholder needs in a timely and useful way and ultimately lead to improvements in human and environmental health; and (2) the Program performance measures and LTGs and priorities of the STS MYP link to desired outcomes and Agency priorities. The Subcommittee also considered the likelihood that the application of products and knowledge by clients would lead to the achievement of program outcomes. Specifically, the Subcommittee answered the following questions:

- How much have the results from P2NT research projects been used by environmental decision-makers to inform decisions and achieve results?

- How well-defined are the Program's measures of outcomes for the STS Research Program?

II.3 Summary Assessment (Program Performance Rated by LTG)

The Subcommittee provided a summary assessment and narrative for LTGs 2 and 3 from the STS MYP. Because no metrics development work was conducted within the P2NT Research Program, no summary assessment was provided for LTG 1. The following summary assessment charge questions were answered by the Subcommittee:

- How appropriate is the science used to achieve each LTG in the STS MYP? (i.e., Is the Program asking the right questions, or has it been eclipsed by advancements in the field?)

By bringing elements such as ecology, land use, water quality, urban planning, and economics into the STS Research Program, a successful transition from the P2NT Research Program to a program on sustainability can be achieved. Otherwise, if those elements are not brought into the Program the science will be left behind and the success of the Program transition will be less likely.

- How good is the scientific quality of the P2NT research products?

- How much have the results from P2NT research products been used by environmental decision-makers to inform decisions and achieve results?

Elements Included for LTG 2

The Subcommittee assessed the appropriateness, quality, and use of P2NT decision support tools and methodologies to assist stakeholders in making decisions and achieving results. The Subcommittee also assessed the extent to which ORD is asking the right questions and conducting the right science to provide tools and methodologies that are responsive to the needs of decision-makers.

Elements Included for LTG 3

The Subcommittee assessed the appropriateness, quality, and utility of P2NT technologies to provide information for solutions to stakeholder issues and environmental problems. The Subcommittee considered the extent to which ORD is asking the right questions and conducting the right science to provide technologies that are responsive to the needs of decision-makers.

Overall Qualitative Score

The BOSC STS Research Subcommittee has assigned a qualitative score that reflects the quality and significance of the research as well as the extent to which the Program is meeting or making measurable progress toward the goal—relative to the evidence provided to the BOSC. The Overall Qualitative Score is given by LTG. The scores are in the form of the adjectives that are defined below and are intended to promote consistency among BOSC program reviews. The adjectives are used as part of a narrative summary of the review, so that the context of the rating and the rationale for selecting a particular rating will be transparent. The rating reflects considerations beyond the *Summary Assessment* questions, and will be explained in the narrative. The adjectives to describe progress are:

- ***Exceptional*** indicates that the program is meeting all and exceeding some of its goals, both in the quality of the science being produced and the speed at which research result tools and methods are being produced. An exceptional rating also indicates that the program is addressing the right questions to achieve its goals. The review should be specific as to which aspects of the program's performance have been exceptional.

- ***Exceeds Expectations*** indicates that the program is meeting all of its goals. It addresses the appropriate scientific questions to meet its goals and the science is competent or better. It exceeds expectations for either the high quality of the science or for the speed at which work products are being produced and milestones met.

- ***Meets Expectations*** indicates that the program is meeting most of its goals. Satisfactory programs live up to expectations in terms of addressing the appropriate scientific questions to meet their goals, and that work products are being produced and milestones are being reached in a timely manner. The quality of the science being done is competent or better.

- ***Not Satisfactory*** indicates that the program is failing to meet a substantial fraction of its goals, or if meeting them, that the achievement of milestones is significantly delayed, or that the questions being addressed are inappropriate or insufficient to meet the intended purpose. Questionable science also is a reason for rating a program as unsatisfactory for a particular LTG. The review should be specific as to which aspects of a program's performance have been inadequate.

III. SCIENCE AND TECHNOLOGY FOR SUSTAINABILITY (STS) RESEARCH PROGRAM

III.1 Program Background

ORD's P2NT Research Program is being combined into a newly organized research program, the STS Research Program, which is focused on the issue of sustainability. Although this represents a new research direction for ORD, the STS Research Program will include a select group of research efforts that had their genesis within the P2NT Research Program. In an effort to demonstrate ORD capabilities, as well as to provide a more complete understanding of the direction the STS Research Program will follow, these P2NT research efforts were presented to the Subcommittee. Key documents that lay the groundwork for this new program are EPA's ORD Sustainability Research Strategy and the STS MYP, both of which were recently reviewed by EPA's SAB.

III.2 Recommendations for the Science and Technology for Sustainability Research Program

III.2.1 Specifics for Program Elements

1. Long-Term Goal 1

The development of sustainability metrics is a critical component of the overall effort, because these are the measures on which the success of all activities needs to be evaluated. It is unclear, however, precisely how the metrics to be developed within this element will be used in other LTGs, and it also is unclear how the metrics to be developed will be informed by activities in the other LTGs.

The current efforts focus on the development of pollution prevention (P2) metrics that are succinct and static measures. An attempt has been made to evaluate P2 technologies using existing metrics. Successful measures have been developed, although it is not clear that EPA has been a leading force in this area. It also is unclear whether the metrics that have been proposed by EPA have permeated into general use. The STS Research Program is working diligently to partner with other institutions, but simply does not have the resources to effectively disseminate their results more widely. The use of endpoints and midpoints is not integrated among the various programs. For example, the ETV Program uses endpoints but the Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts (TRACI) is based on midpoints. Some metrics of performance are excellent. For instance, in the ETV Program, metrics used to evaluate the performance of technologies have been successful. The ETV Program evaluates devices that measure important sustainability metrics that would not be done elsewhere.

The Technology for a Sustainable Environment (TSE) Program was a very strong program with innovation, productivity, and highly cited papers. It was a cost-effective way to enhance participation in these research questions. An extramural program could be crafted to emphasize metrics and how technologies move towards improving the measures.

This program element faces a difficult challenge—that of converting very complex concepts into simply understandable measures. The use of rigorous metrics is critically important in the development of decision-making tools and also should drive research needs in both internal and external programs. A key component of the development and testing of appropriate metrics is a clear conceptual definition of what is to be measured with a particular set of metrics. Thus, clear definitions of the sustainability concepts being addressed and the component elements of these concepts are required before a specific metric can be assessed and its applicability in real-world situations evaluated. Additional attention needs to be given in this element to the process of development and evaluation of sustainability metrics. Testing of specific metrics in real-world situations also is appropriate, but one needs to propose and develop the metrics first. Then, the testing protocols should be established to determine if the metrics are measuring the intended functions, if they are consistent in their evaluation, if they are sufficiently independent, and if they can be effectively used to determine if specific actions are driving society to become more sustainable.

2. Long-Term Goal 2

Within the P2NT Research Program the development of LCA is highly relevant and important to the goals of the EPA, stakeholders, and the international community. The tools developed by EPA already have been demonstrated to be successfully delivered under the P2NT strategy and MYP. The decision-making tools being developed are well aligned with the overall Program's priorities and goals. LCA is the method of choice for examining the overall inputs and outputs of a manufacturing or other process. For example, TRACI is routinely used by academic and industrial stakeholders across the globe as a way to evaluate environmental life cycle impacts. As the Program morphs to a sustainability-oriented decision-making process, the LCA aspect will become critical. There are two other critical aspects under development by the Program that will support the use of LCA. First, there is a project underway to better understand the environmental effects of different processes. The relationship between process outputs and specific environmental impacts is still an important research area in LCA. A second aspect of the proposed research plan that will be critical is how to incorporate spatial and temporal relationships. An industrial process does not occur in isolation, but is dependent on the surrounding economic and environmental landscape. The industrial process also will occur over a timeframe of decades, which will see changes in the surrounding landscapes. It will be advantageous to have decision-making tools that can incorporate time and space in a native fashion.

3. Long-Term Goal 3

The P3, SBIR, and ETV Programs are highly relevant from a P2NT perspective. There is evidence of public benefit, from the meeting of program goals in all three programs. The Green Technology Program is less relevant and there is a less clear connection to Program goals. It was

not clear what activities were underway that could not be accomplished in the public or academic sectors in the absence of the EPA program; thus, the public benefits were not evident. There seems to be no integrated plan to meet Agency/client needs. Rather, the research areas seem to be more driven by investigator interests.

Relative to the STS transition, the P3, SBIR, and ETV Programs remain integral. These three programs are basically solid, but with additional resources could further increase their relevance. The solicitation/judging criteria for the P3 Program should be improved to require a clear statement by students as to effects articulated via sustainability metrics or decision tools. More emphasis should be placed on measurement. This will force students to more clearly articulate how their projects relate to sustainability. The ETV Program should encourage an increased role in supporting emerging markets in trades/mitigation/offsets, such as mercury/greenhouse gases, etc. An analysis should be conducted to determine if there are emerging markets in the trade/offset areas that have barriers to verification. Then, research could focus on how to solve or minimize these impediments to verification and subsequent use. The SBIR Program should increase the use of sustainability metrics in the selection criteria. It also should increase the linkages between Program outcomes and sustainability metrics.

The Green Technology Program as currently configured might be perceived to be largely irrelevant. Consideration should be given to redirecting the Program or replacing it with an extramural grants program. The Subcommittee had concerns about the selection criteria for efforts. The Program should consider what the marketplace will do well and what will not be done well and then choose to influence those sectors where there is an impediment, primarily a lack of information or structure. The Program then should select research needs based on what is needed that the marketplace is not well structured to meet. The Program should assess whether there are key components of market failures and/or Agency goals that would benefit from a targeted internal program, such as lead poisoning prevention in environmental justice communities or arsenic removal on tribal lands. Even if these types of assessments are conducted, the Subcommittee is unsure of the resulting public benefits. Again, many of these needs can be met by local jurisdictions after a framework has been established. The Program could use an extramural funding approach instead of in-house capability. The Program is encouraged to use an LCA framework or approach to identify market failure points in a broader process than currently is being addressed in the private sector. Then, the Program could leverage more outcomes from Agency-sponsored activities on a process with potential widespread commercial application with high marginal benefit.

III.2.2 Program Structure

A. Comments on Charge Questions

In general, the Program structure appears to be adequate but it should be assured that there is integration and continuity among the elements during the plan for transition. The existing program elements and the structure proposed for STS research in the future are organized around the development of scientifically-based sustainability metrics. The current structure of the Program and the proposed structure are well-suited for the development of decision support tools that promote environmental stewardship and sustainable management practices. The Program as

planned will be able to develop, apply, and demonstrate innovative technologies that solve environmental problems and provide sustainable outcomes.

B. Specifics for Program Elements

1. Long-Term Goal 1

Sustainability targets need to be identified so that appropriate metrics can be designed and tested. It was unclear if the measurements that are being made validate the metrics that are being defined or if there were testable hypotheses included. Critical experiments should be designed that allow testing of hypotheses within the realm of defined metrics. The predictability of the models should be evaluated and sensitivity analyses conducted. Evaluation of the metrics should be done systematically and quantitatively. Evaluative criteria should be developed and tested. The two Annual Performance Goals (APGs) do not seem to flow well into a logical research plan, with quantifiable goals and objectives.

2. Long-Term Goal 2

The STS Research Program has not yet been completely developed, thus not enough information is available at this time to comment on the Program's structure. From the information available, the proposed STS Research Program is logically organized and the proposed organizational structure makes maximum use of the personnel available. The lines of communication are short and well-defined and the communication among subgroups appears to be adequate. Definitions are needed for some terms to improve clarity of Program elements and responsibilities. It is suggested that some additional elements or aspects be implemented.

3. Long-Term Goal 3

The science being conducted by the STS Research Program is clearly appropriate and has historically made significant contributions to the development and demonstration of the utility of innovative technologies that solve environmental problems and provide sustainable outcomes. This has been particularly true in the past. Much of the work currently being conducted by the STS Research Program, however, is eclipsed by the magnitude and pace of advancements of industrial and academic communities.

The STS MYP flow of work reasonably reflects the anticipated pace of scientific progress and timing of client needs. The STS MYP indicates a reasonable sequence of events. It is sparsely populated, however, and not coordinated with outside efforts. The STS Research Program design is logical, but the transition from P2NT seems less well planned. Each individual appears to view his or her own work as the highest priority. To an outside reader, the STS Research Program priorities are clearly identified.

III.2.3 Program Quality

Specifics for Program Elements

1. Long-Term Goal 1

The STS Research Program has had a large impact by identifying and creating scientifically-based sustainability metrics. Some of the algorithms developed have been implemented into standard process simulators. Within the P2NT Research Program, appropriate scientific elements were generally included and the quality was high. It is unclear, however, that access to additional sustainability elements will be present as the Program transitions, given the limited staff and interactions with other programs at EPA. It also was unclear if this LTG has been integrated into a competitive extramural funding program. This might be a mechanism to further leverage funds to have a greater impact in this area.

2. Long-Term Goal 2

The tools developed under the P2NT Research Program have been of high scientific quality and were developed using appropriate and current science. It would be highly beneficial to have collaborations with both industry and academia to capture the needs of the community and to be able to develop and/or promote the appropriate tools with scientific basis. These tools could be U.S.-focused databases, models for estimations, etc. The overall quality of the Program, relative to addressing LTG 2, could be improved through targeted extramural collaborations on the development of new tools or cooperation on the advancement of existing tools or tools being developed in the private sector.

Currently, it does not appear as if extramural collaborations are planned on techniques to better relate process outputs to environmental impacts. In the future, impacts will include outcomes beyond simple toxicity, but may include changes in habitat, added nutrients, or even the increased urbanization of an area due to an increase in employment opportunities. These are all areas of research beyond the available manpower of the current program. Research into understanding spatial and temporal relationships of an activity is a developing field for which the current staff has little expertise. Interaction with extramural scientists has the potential to rapidly add these capabilities to the toolbox.

3. Long-Term Goal 3

The scientific quality of the Program is very high and representative of a collection of individual experts in specific topic areas of development, application, and demonstration of innovative technologies that solve environmental problems and provide sustainable outcomes. The related communications are well-written, but apparently less well- (and widely) read, in part because the general state-of-the-science may have surpassed the timeline of the P2NT Research Program.

The science used in the P2NT research is appropriate, although, in some cases, untimely. By the time it is "perfected" and published, the outside user community has moved on. In other cases, the product is overly complicated and not usable by the desired clientele. For example, it is of

the utmost importance in the industrial sectors to have tools for streamlining LCAs that allow for rapid evaluation of environmental burdens. Several companies have moved beyond tools developed by the Agency and have developed their own, as the timeframe for action in the industrial sector is significantly shorter.

The quality of the research conducted under the P2NT Research Program is high. The staff understands the peer review process and has an understanding of expected standards. The extent to which competitive funding affects the overall quality of the products is less clear. Appropriate means have been employed to ensure quality, including peer review and competitive funding.

III.2.4 Program Leadership

Specifics for Program Elements

1. Long-Term Goal 1

Given their limited numbers, the team leaders are having an appropriate impact on the development of scientifically-based sustainability metrics. A team that was better integrated throughout EPA could draw on additional resources that could enhance its effectiveness.

2. Long-Term Goal 2

The scientists clearly are recognized by the international scientific community as leaders in the development of decision-support tools that promote environmental stewardship and sustainable management practices, such as the LCA Program in EPA. One of the clear examples is the reputation that the LCA Program has developed in North America and Europe. LCA now is an integral part of the European decision-making process, led in part by the EPA effort. The LCA research also has been extensively cited in the peer-reviewed literature. The new initiatives that have been proposed to develop additional sets of tools are too early in the developmental cycle to evaluate. A word of caution, however, in that continued support of this area is important to keep abreast of and continue to lead the development of LCA methodologies.

3. Long-Term Goal 3

Until more information is generated by the individuals within the P3 Program it is impossible to make an assessment. Thus, a detailed analysis of the impacts on the P3 Program on the student participants is desirable. Similarly, for the ETV Program, it is not possible to assess how the program has impacted the careers of participants. The ETV Program is a very credible program with a highly developed degree of leadership/quality. The Green Technology Program, at the time of initiation, was clearly of high quality and made a significant impact on the development and application of innovative technologies that solve environmental problems and provide sustainable outcomes. The initial work of this team was of very high quality and the researchers were recognized as world leaders in this area. It is less clear if this continues to be the case. In fact, the limited resources available to this group virtually assures that their efforts will be eclipsed by others as the Program proceeds.

III.2.5 Coordination and Communication

Specifics for Program Elements

1. Long-Term Goal 1

There needs to be significant interaction between this LTG and others, particularly LTG 1 and LTG 2, which are intimately tied together. In addition, LTG 1 metrics should be used to inform LTG 3 activities.

2. Long-Term Goal 2

The P2NT Research Program is highly coordinated within the Agency and there is evidence of good coordination and interaction outside the Agency, especially with professional organizations such as the International Organization for Standardization (ISO), the Society of Environmental Toxicology and Chemistry (SETAC), and others. There also is good communication flow and exchange of ideas with some industrial sectors, a practice that should, whenever possible, be encouraged. Many of the thought leaders are to be found in that sector and EPA cannot bring the same level of resources to bear on the issues as can the private sector. It is not clear if the coordination has been successful in reaching a wider set of stakeholders, such as NGOs, state agencies, etc. If possible, these efforts should be encouraged. The research results have been communicated internally and externally, but the actual outputs and outcomes could be defined more clearly and more effectively communicated to targeted sectors.

3. Long-Term Goal 3

Many of the results of the research under LTG 3 have not reached the user community. Some outside partners appear dedicated to obtaining grant support and/or license fees as opposed to utilization of the technology. Also, there is a need to better understand what has been done by academia and industry.

While it is clear that ORD has collaborated with and obtained input from others on research objectives, especially to avoid duplication of effort, the Subcommittee members thought that this is so critical to the acceptance and use of the technologies developed that the Program should seek input from a number of extramural groups to assist EPA. In fact, such communications could result in partnerships and greater leveraging of the limited Program resources. Some of the work is a duplication of previous or current work being done by others outside of EPA. Larger industrial and manufacturing firms are underrepresented in setting objectives and avoiding duplication. Hurdles should be lowered and/or obstacles removed to insure that "those who really know" participate. It appears that the U.S. Department of Energy (DOE) and others may be acting without knowledge of the P2NT efforts. While not a complete replication, the projects and programs seem to be uncoordinated with those of other agencies.

The results of the work have not been effectively communicated to larger industrial enterprises. Booths at meetings and brochures with "one-pagers" are far less effective in conveying such messages than are presentations, publications, and patents. Internally, it appears old "stove-

pipes" still hinder the transfer of information. Concentrating on having efforts "subject to easy discovery" on the Internet also would be beneficial.

III.2.6 Outcomes

EPA must perform scientific research to provide the baseline knowledge to establish technical standards for regulatory decisions and to provide appropriate tools and methodologies to those who are charged with measuring and conserving the health of our national environment. In this section, the Subcommittee asks whether the STS Research Program LTGs are appropriate and well integrated and whether the Program is effective in meeting its responsibilities to its customers. To the extent possible, the Subcommittee also assesses future directions of STS research. In evaluating the Program and its LTGs, it is important to reiterate that ORD has a research mission with the primary responsibility of developing scientific knowledge, tools, and methodologies to serve the mission of the regulatory divisions of EPA, as well as to meet the technical and scientific needs of state governments and tribes. In this sense, other divisions of EPA are the customers of ORD and its constituent research programs. Other important customers include states, tribal governments, and local decision-makers.

Specifics for Program Elements

1. Long-Term Goal 1

The metrics developed under the P2NT Research Program have not pervaded other programs. The APGs should be provided in more quantifiable forms, generally in the form of SMART (specific, measurable, achievable, relevant, and timely) goals. The goals are written very generically, without sufficient measurable targets against which one can evaluate performance. APM 1, 2008 is well-defined, but 2009 is nebulous and could be refined.

2. Long-Term Goal 2

The Subcommittee determined that the STS Research Program was exceeding expectations relative to LTG 2. The Program is relatively mature in this area and a great deal of progress has been made. The progress toward achieving this LTG has been excellent and has had a large impact on the field of sustainability. Because of the excellence of this Program and the impact that it has had on the field of sustainability, the Subcommittee felt compelled to highlight several accomplishments within this program element.

One example of a program with many successful elements is the ETV Program. The ETV Program develops testing protocols and verifies the performance of innovative technologies that have the potential to improve protection of human health and the environment. The ETV Program was created to accelerate the entrance of new environmental technologies into the domestic and international marketplaces. The Subcommittee would like to recognize two program elements that it considered to be of excellence. These include: (1) the public outreach component which brings early public use; and (2) the clear team spirit of the Program members. To find a balance of speed and a team sense of "over-accomplishment" is rare. ORD can be rightly proud of this program element and the impact that it has had. The Subcommittee recommends this program element for an ORD citation if this has not been done already.

Specific outstanding contributions were noted by the Subcommittee in the area of metrics/tools. This is not meant to be to the exclusion of other excellent program elements, but to be a concrete example of excellence to which other program elements could aspire. The development of TRACI and BEES (Building for Environmental and Economic Sustainability), which is a trademark registered with the U.S. Patent and Trademark Office (USPTO), represent two outstanding accomplishments that have "driven" the science and resulted in tools that are used and have had a large impact. Thus, the Subcommittee would suggest that these accomplishments, while representing "outputs," also can be characterized as "outcomes." The first is the TRACI tool, which is recognized throughout the world as a comprehensive means of tracking emissions through to ultimate effects while providing midpoint indicators. The latter point bridges the gap between "process indicators" and "outcome indicators." The BEES tool offers other lessons of value. First, it has a progression of versions (3.0c during the BOSC meeting; 4.0 today). The Subcommittee was pleased to see that once a viable tool was developed, it was continually updated to keep it relevant with the changing science and global conditions. The second key point is that BEES is a model of interagency cooperation, between EPA and the National Institute of Standards and Technology (NIST). The BEES model is widely used by a variety of government agencies and voluntarily by industry for green purchasing.

3. Long-Term Goal 3

The Program could benefit from a more systematic evaluation of the Program outcomes. Tracking of all participants (not just the winners) could provide useful information for future planning. Once the participants move to a more careful statement of sustainability, effects in terms of some metric(s) then can track outcomes in these terms from the implementation phase (if any).

ETV: The current outcomes analysis does not measure the effect of the ETV Program, because it does not attempt to identify outcomes in the absence of the ETV Program. Outcome measures stated in terms of numbers of verifications are probably better, even if considered less relevant. These metrics are better linked to the question of number of decision-makers/impacts. The metrics were deemed to be well-defined, but not well-quantified.

SBIR: Increase meeting of stakeholder needs. If the Program can better address the internal Agency needs from the STS MYP it will provide a valuable service and be recognized more favorably. The goal of moving to a 100 percent cost share basis needs to be carefully evaluated. Although this will better leverage funds, it might miss important opportunities. This might be a future goal, but it needs to be determined if this would result in missed opportunities for small businesses. This could occur if they: (1) could not afford the assessment; and (2) are not being funded for this purpose through the SBIR Program. Additional SBIR opportunities in the broader set of sustainability concerns, such as land and water uses, need to be explored. One example might be the design of storm water handling systems in new developments. Certainly there are other opportunities as well.

Green Technology: The Program's outcomes are not defined in terms of the STS Research Program goals. They are statements of effects, not measures of outcomes. The term "better use

of resources" does not demonstrate a reduction (on a total cost accounting basis) from cradle-to-grave perspective. Similarly, the statement "promoting the quality of life…." is not sufficiently descriptive.

III.2.7 General Recommendations

Long-Term Goal 1

The Program has just begun and there has not been enough time to make a long-term recommendation.

Long-Term Goal 2

Overall Assessment

Overall, the Program was considered to be well thought out and designed, with strong alignment to general sustainability concepts. For instance, the LCA Program is of world leadership, well-developed, widely used, and recognized by stakeholders. Development of streamlined methods is needed as part of the expansion of LCA tools (e.g., make them user-friendly) as well as integration of material flow analysis (e.g., industrial ecology concepts). The MYP goal of creating sustainable management decisions on future outcomes for use at the local, regional, national, and international levels is a strong part of the Program. The inclusion of knowledge transfer from lessons learned from collaborative projects that addressed sustainable solutions is a strong aspect of the MYP, and needs to be tied to a communication goal. System-based methods are indispensable for moving towards sustainability. These are integrated in the STS MYP, but need to be integrated into tools.

Specific Recommendations

The Subcommittee recommends integrating an implementation plan as part of the STS MYP. Some concepts in the APGs of the STS MYP need to be defined (e.g., 'sustainable land use,' 'sustainable water use,' 'local level') to ensure clear understanding by stakeholders and to ensure that all the aspects of sustainability are incorporated. Strengthen and expand communication aspects of tools as part of the MYP including: (1) guidance regarding scope (e.g., what LCA does and does not do), outreach, and influence (how LTG 1, LTG 2, and LTG 3 tie together in the path to sustainability); and (2) interrelations of different aspects of sustainability. Ecological aspects should be incorporated into the decision analysis tools. Additional expertise might be needed to cover ecological systems so it would be wise to strengthen collaborations with the ORD Ecology Research Program. Geographic and landscape orientation should be incorporated for local implementation. Are temporal and spatial aspects included? An example of this is the Hierarchical Patch Dynamics Paradigm (HPDP) (Wu and David, 2002). Data transparency, sensitivity assessment, and uncertainty analysis should be routinely included in the tools (e.g., Are numbers truly significant to two decimal places?). The Program should incorporate additional decision-making tools, such as probabilistic risk assessment, Bayesian networks, causal pathways, and Multi-Criteria Decision Analysis (Igor Linkov and others) in the research program. Additional stressors beyond chemicals (e.g., invasive species, water allocation) also should be incorporated. Extramural programs such as the Collaborative Science and Technology Network for Sustainability (CNS) Program should be incorporated and/or continued. Economics

and other social dimensions should be incorporated as part of feedback loops of process or output evaluated decision-making.

Long-Term Goal 3

Overall Assessment

The overall assessment of LTG 3 is that many components of the programs are of high quality, and that generally, the programs under this goal are relevant and well-structured, and use the appropriate science. In general, the Program is using the appropriate science to address the appropriate questions, in a manner that is being used by decision-makers. The impact of the Program will continue to be limited by resources to the point that a "critical mass" may not be possible to achieve or sustain. In general, the program elements have components that would achieve the STS MYP elements. The transition to a full sustainability program, however, will require a clearer linkage to the other LTGs, especially the metrics being used, and a broader view of client needs and Agency goals as defined by the six themes of sustainability.

Specific Recommendations

All of the program elements and the Green Technology Program in particular are in need of refinement to better address sustainability issues and to demonstrate and articulate the role that they play in contributing to sustainable outcomes. Some specific suggestions by program element are given below.

P3: Integrate sustainability metrics into judging criteria.

SBIR: Integrate potential impact on sustainability metrics into program solicitations and selections, and into program evaluation.

ETV: Broaden the mission to evaluate and verify additional components of the sustainability program and look for opportunities to support emerging markets in trading, offsets, and mitigation.

Green Technology: Carefully examine the rationale for the selection of target areas/technologies to better address market failures and tie outcome measures to sustainable measures and metrics.

III.2.8 Recommendations Relative to Summary Assessment

Overall STS Research Program

Qualitative Score: *Meets Expectations*

The STS Research Subcommittee that conducted the review found that the STS Research Program is Meeting Expectations (see section II.3 for an explanation of the meaning of this qualitative score and other possible scores). Some elements of the Program are excellent and exceed expectations. Where the Program does not exceed expectations the primary reason is that these program elements are small and lack critical mass. Other elements were difficult to judge because they are in transition. The STS Research Program has some excellent researchers who

are world leaders in their fields. The quality of the research is apparent. The Program is doing much with relatively limited resources. In particular, the leveraging of available resources by partnering with other agencies, both local and federal, allows the STS Research Program to achieve more than it otherwise could. The limited resources, however, direct the types of studies that are undertaken. For this reason, the research undertaken might not include the greatest priorities of the Program, or may not move the overall science forward as rapidly as otherwise could be achieved. This resource-driven approach dictates that the STS Research Program will not be the leader that it might otherwise have the potential to be.

- ✍ It should be assured that there is integration and continuity among the elements during the plan for transition.

- ✍ The potential impact of STS programs is limited by lack of a critical mass and resources. In developing the STS Research Program, ORD must make better use of capabilities across ORD.

- ✍ Currently, much of the work being conducted by the STS Research Program is eclipsed by the magnitude and pace of advancements of industrial and academic communities. Thus, in developing the plan, the Program must make strategic decisions on where it can make an impact on the overall field.

- ✍ Because the STS Research Program is sparsely populated and not coordinated with outside efforts, a strategic plan that includes an awareness of what is being done outside of the Agency, including that of organizations outside of the United States, and how ORD can make a significant impact on the science should be developed.

1. Long-Term Goal 1

Qualitative Score: ***No Score Given***

Because LTG 1 only now is being developed and the Program reorganized, there was insufficient information to allow the Subcommittee to give a qualitative score. The Subcommittee reviewed the plan, however, and provides some overall guidance below:

- ✍ Develop an outline for how metrics for sustainability will be developed. This should include criteria for assessing the utility and predictability of metrics.

- ✍ Coordinate metric development with other LTGs.

- ✍ Determine a strategy of how metrics will be used.

- ✍ Going forward, an extramural program based on the TSE Program could be crafted to emphasize metrics and how technologies move toward improving the measures.

- ✍ Testing protocols should be established to determine if the metrics are measuring the intended functions, if they are consistent in their evaluation, if they are sufficiently

independent, and if they can be effectively used to determine if specific actions are driving society to become more sustainable.

- Sustainability targets need to be identified so that appropriate metrics can be designed and tested.

- Critical experiments should be designed that allow testing of hypotheses within the realm of defined metrics.

- The predictability of the models should be evaluated and sensitivity analyses conducted.

- Evaluation of the metrics should be done systematically and quantitatively.

- The team should be better integrated throughout EPA in order to draw in additional resources that could enhance its effectiveness.

- There needs to be significant interaction between this LTG and others, particularly LTG 1 and LTG 2, which are intimately tied together.

- LTG 1 metrics should be used to inform LTG 3 activities.

2. Long-Term Goal 2

Qualitative Score: ***Exceeds Expectations***

The Subcommittee determined that the STS Research Program was exceeding expectations relative to LTG 2. The Program is relatively mature in this area and a great deal of progress has been made. The progress toward achieving this LTG has been excellent and has had a large impact on the field of sustainability. The Subcommittee provides a few observations and suggestions below:

- The LCA programs, metrics, and procedures developed under the P2NT Research Program are relevant and important to the goals of the EPA, stakeholders, and in the international community. The STS Research Program is positioned to move these initiatives forward and is encouraged to build on this strength.

- LTG 2 could be improved through targeted extramural collaborations on the development of new tools or cooperation on the advancement of existing tools or tools being developed in the private sector.

- Efforts should be made to reach a wider set of stakeholders, such as NGOs, state agencies, etc.

- The actual outputs and outcomes could be more clearly defined and communicated to targeted sectors.

3. Long-Term Goal 3

Qualitative Score: ***Meets Expectations***

Although the Subcommittee found the overall performance of the STS Research Program, relative to LTG 3, to be meeting expectations, a range of performances were observed. Some program elements were performing at a very high level and some would be classified as exceeding expectations. As a whole, however, the Subcommittee members thought that the overall performance was "meeting expectations." Some observations and recommendations are given below:

- The P3, SBIR, and ETV Programs have all been highly relevant to the mission of the EPA and the program elements in these programs should be preserved whenever possible.

- The relevance and impact of the Green Technology Program is less apparent and this program needs to be assessed internally to determine if it is serving a function that is not being met already by the private sector and academia.

- The solicitation/judging criteria for the P3 Program should be improved to require a clear statement by students as to effects articulated via sustainability metrics or decision tools.

- More emphasis should be placed on measurement and quantitative assessment of outcomes within the P3 Program.

- The P3 Program could benefit from a more systematic evaluation of the program outcomes, such as tracking of careers of recipients to obtain information that can be used to measure impact as outcome.

- The ETV Program should encourage an increased role in supporting emerging issues in trades/mitigation/offsets, such as mercury/greenhouse gases, etc.

- An analysis should be conducted to determine if there are emerging markets in this trade/offset area that have a barrier surrounding verification issues.

- The SBIR Program should increase its use of sustainability metrics in selection criteria and increase the linkage of program outcomes to sustainability metrics.

- Consideration should be given to redirecting the Green Technology Program or replacing it with an extramural grants program.

- Results from the Green Technology Program have not been effectively communicated to larger industrial enterprises.

- The Program could benefit from a more systematic evaluation of the program outcomes, such as tracking of careers of recipients to obtain information that can be used to measure impact as outcome.

Acknowledgements

The Subcommittee would like to thank the administrators and staff for their assistance with this program review. The STS Research Program provided all of the materials requested by the Subcommittee. The Subcommittee appreciated the timely and very professional manner in which all of the EPA administration and staff provided materials, both written and oral. It was evident that a great deal of time and thought had gone into the preparations for the program review. The Subcommittee recognizes that this is time diverted from other activities and appreciates the excellent cooperation it received from all of the STS Research Program Managers. It also was clear that the STS Research Program has very dedicated staff members, some of whom are world leaders in their fields. The observations and recommendations made by the Subcommittee are submitted in a spirit of collegiality in the hope that they will assist the STS Research Program during its transition and allow it to be as effective as possible.

IV. REFERENCES

[1] Budget Data Request 04-31. Executive Office of the President, Office of Management and Budget. March 22, 2004. "Completing the Program Assessment Rating Tool (PART) for the FY06 Review Process," pp. 50-56.

[2] Memorandum for the Heads of Executive Departments and Agencies. Executive Office of the President, Office of Management and Budget. June 5, 2003. "FY 2005 Interagency Research and Development Priorities," pp. 5-10.

[3] Evaluating Federal Research under the Government Performance and Results Act. (National Research Council, 1999).

[4] The House Science Subcommittee. Letter to Dr. Bruce Alberts, President of the National Academy of Sciences, from F. James Sensenbrenner, Jr. and George E. Brown. October 23, 1997.

[5] The Government Performance and Results Act: 1997 Government-Wide Implementation Will Be Uneven. U.S. General Accounting Office (GAO/GGD, 1997).

[6] Building a Foundation for Sound Environmental Decisions. (National Research Council, 1997).

[7] "Renewing the Compact between Science and Government," Stokes, D.E. In: 1995 Forum Proceedings, Vannevar Bush II—Science for the 21st Century. Research Triangle Park, NC: Sigma Xi, 1995, pp. 15-32.

[8] Risk Assessment in the Federal Government: Managing the Process. (National Research Council, 1983).

[9] Strengthening Science at the U.S. Environmental Protection Agency. (National Research Council, 2000 p. 141).

V. APPENDICES

Appendix A: Science and Technology for Sustainability Subcommittee Members

Chair

John P. Giesy, Ph.D.
Professor and Canada Research Chair in Environmental Toxicology
Department of Veterinary Biomedical Sciences and Toxicology Centre
University of Saskatchewan
44 Campus Drive
Saskatoon SK S7N 5B3
Canada
Phone: 306-966-7441
Fax: 306-931-1664
E-mail: jgiesy@aol.com

Subcommittee Members

Wayne Landis, Ph.D.
Director, Institute of Environmental Toxicology
Chair, Department of Environmental Sciences
Huxley College of the Environment
Western Washington University
Bellingham, WA 98225-9180
Phone: 360-650-6136
Fax: 360-650-6556
E-mail: wayne.landis@wwu.edu

Concepción Jiménez-González, Ph.D.
GlaxoSmithKline, Corporate EHS
5 Moore Drive
Research Triangle Park, NC 27709-3398
Phone : 919-315-2139
E-mail: conchita.j.gonzalez@gsk.com

Earl R. Beaver, Ph.D., FAIChE
Chair Emeritus
Institute for Sustainability
14500 White Birch Valley Lane
Chesterfield, MO 63017-2418
Phone: 636-536-1256
E-mail: erbeav@aol.com

Martin A. Abraham, Ph.D.
Dean, College of Science, Technology, Engineering and Mathematics
Youngstown State University
Youngstown, Ohio 44555
Phone: 330-941-3009
E-mail: martin.abraham@ysu.edu

Ted Tomasi, Ph.D.
Vice President and Technical Director
ENTRIX, Inc.
10 Corporate Circle, Suite 300
New Castle, DE 19720
Phone: 302-395-1919
Fax: 302-395-1920
E-mail: ttomasi@entrix.com

Consultant to the Subcommittee

Peter Blaze Corcoran, Ed.D.
Professor of Environmental Studies and Environmental Education
Director of the Center for Environmental and Sustainability Education
College of Arts and Sciences
Florida Gulf Coast University
10501 FGCU Boulevard South
Fort Myers, FL 33965-6565
Phone: 239-590-7166
Fax: 239-590-7200
E-mail: pcorcora@fgcu.edu

Appendix B: List of Science and Technology for Sustainability Subcommittee Meetings

January 23, 2007	Conference Call
February 21, 2007	Conference Call
April 25–26, 2007	Site Visit at EPA Facilities, Cincinnati, Ohio
May 30, 2007	Conference Call
September 6, 2007	Conference Call
December 11, 2007	Conference Call

Appendix C: Science and Technology for Sustainability Subcommittee Draft Charge

DRAFT Program Review Charge
Science and Technology for Sustainability Subcommittee

1.0 Objective. The BOSC Science and Technology for Sustainability Subcommittee will conduct a prospective and retrospective review of ORD's Science and Technology for Sustainability Research Program, and evaluate the Program's relevance, quality, performance, and scientific leadership. The BOSC's evaluation and recommendations will provide guidance to ORD to help:

- plan, implement, and strengthen the program;
- compare the program with programs designed to achieve similar outcomes in other parts of EPA and in other federal agencies;
- make research investment decisions over the next 5 years;
- prepare EPA's performance and accountability reports to Congress under the Government Performance and Results Act; and
- respond to assessments of federal research programs such as those conducted by the Office of Management and Budget (OMB highlights the value of recommendations from independent expert panels in guidance to federal agencies.[1,2]).

2.0 Background Information. Independent expert review is used extensively in industry, federal agencies, Congressional committees, and academia. The National Academy of Sciences has recommended this approach for evaluating federal research programs.[3]

Because of the nature of research, it is not possible to measure the creation of new knowledge as it develops—or the pace at which research progresses or scientific breakthroughs occur. Demonstrating research contributions to outcomes is very challenging[4] when federal agencies conduct research to support regulatory decisions, and then rely on third parties[5]–such as state environmental agencies—to enforce the regulations and demonstrate environmental improvements. Typically, many years may be required for practical research applications to be developed and decades may be required for some research outcomes to be achieved.

Most of ORD's environmental research programs investigate complex environmental problems and processes, combining use-inspired basic research[6,7] with applied research, and integrating several scientific disciplines across a conceptual framework[8] that links research to environmental decisions or environmental outcomes. In multidisciplinary research programs such as these, progress toward outcomes can not be measured by outputs created in a single year. Rather, research progress occurs over several years, as research teams explore hypotheses with individual studies, interpret research findings, and then develop hypotheses for future studies.

In designing and managing its research programs, ORD emphasizes the importance of identifying priority research questions or topics to guide the research. Similarly, ORD recommends that its programs develop a small number of performance goals which serve as indicators of progress to answer the priority questions and to accomplish outcomes. Short-term outcomes are accomplished when research is applied by specific clients—to strengthen environmental decisions or regulations, for example. These decisions and resulting actions (e.g., the reduction of contaminant emissions or restoration of ecosystems) ultimately contribute to improved environmental quality and health.

In a comprehensive evaluation of science and research at EPA, the National Research Council[9] recommended that the Agency substantially increase its efforts to explain the significance of its research products and to assist clients inside and outside the Agency in applying them. In response to this recommendation, ORD has engaged science advisors from client organizations to serve as members of its research program teams. These teams help identify research contributions with significant decision-making value and help plan for their transfer and application.

For ORD's environmental research programs, periodic retrospective analysis at intervals of 4 or 5 years is needed to characterize research progress, to identify when clients are applying research to strengthen environmental decisions, and to evaluate client feedback about the research. Conducting program evaluation at this interval enables assessment of research progress, the scientific quality and decision-making value of the research, and whether research progress has resulted in short-term outcomes for specific clients.

A description of the OSTP/OMB *Research and Development Investment Criteria* is included in Appendix 1.

3.0 Background for the Science and Technology for Sustainability Research Program and Draft Charge Questions. As ORD's Pollution Prevention and New Technologies (P2NT) Research Program is nearing its completion, a new research program has been created, the Science and Technology for Sustainability (STS) Research Program, that is focused on the question of sustainability. While this represents a new research direction for ORD, the STS Research Program will include a select group of research efforts that had their genesis within the P2NT Research Program. In an effort to demonstrate ORD capabilities, as well as to provide a more complete understanding of the direction the STS Research Program will follow, these P2NT research efforts will be presented to the Subcommittee. Key documents that lay the groundwork for this new program are EPA's ORD Sustainability Research Strategy and the STS Multi-Year Plan (MYP), both of which were recently reviewed by EPA's Science Advisory Board.

(A) Program Assessment (evaluate entire research program): The responses to the program assessment charge questions listed below should be in a narrative format, and should capture the performance for those aspects of either the P2NT or STS Research Programs (as identified), or for all activities supporting the Program's Long Term Goals (LTGs).

Program Relevance

- How relevant and consistent has P2NT research been with respect to Agency goals and customer needs?
- How evident are the public benefits of the P2NT research?
- How consistent are the LTGs of the STS Research Program with achieving the Agency's strategic plan?
- How responsive is the new STS Research Program direction to client needs and recommendations from outside advisory boards?

Factors to consider**:** Whether P2NT outputs were/are used by the Agency and stakeholders; the STS Research Program addresses issues raised within the Sustainability Research Strategy; the STS MYP has clear goals and priorities; stakeholders (e.g., program and regional offices) are involved in the planning and prioritization of the research; outputs from the STS Research Program are likely to be used by stakeholders; and the STS Research Program outlines a well-coordinated effort with outside research organizations, nationally and internationally, that will avoid duplication of effort and promote synergistic collaboration.

Program Structure

- How clear and logical are the LTGs in the STS MYP for organizing and planning the research and demonstrating outcomes of the Program?
- How appropriate is the science used to achieve each LTG in the STS MYP (i.e., Is the Program asking the right questions, or has it been eclipsed by advancements in the field?)?
- To what extent does the STS MYP describe an appropriate flow of work that reasonably reflects the anticipated pace of scientific progress and timing of client needs?
- How logical is the STS Research Program design? How clearly identified are the STS Research Program priorities?

Factors to consider: Whether the STS MYP addresses research questions identified in the Sustainability Research Strategy, uses appropriate criteria to select research projects, and if the LTGs are appropriate for planning the research and for identifying long-term priorities that meet the scientific needs of the Agency and Program customers.

Program Quality

- How good is the scientific quality of the P2NT research products?
- How appropriate is the science that has been used in the P2NT research?
- To what extent have appropriate means been employed to ensure quality P2NT research (including peer review, competitive funding, etc.)?

Factors to consider: Whether the research involved a sufficient number of internal and external peer reviews and was modified in response to these reviews; used state-of-the art

science; and used a competitive merit based process that maintained quality when awarding extramural funds.

Scientific Leadership

? Please comment on the leadership role that ORD staff has had in contributing to advancing the current state-of-the-science for tools, methodologies, and technologies that support environmental decision-making.

Factors to consider: The degree to which ORD staff are seen as leaders in the field and active participants in national and international science and technology professional bodies.

Coordination and Communication

? How effectively has ORD engaged outside organizations (both within and outside government) for the P2NT research? How effectively has ORD collaborated with and obtained input from others on research objectives, especially to avoid duplication of effort?
? How effective were the mechanisms used for communicating research results for the P2NT Research Program, both internally and externally?

Factors to consider: The extent to which the Program collaborated with clients and stakeholders and leveraged resources from related research programs.

Outcomes

? How much have the results from P2NT research projects been used by environmental decision-makers to inform decisions and achieve results?
? How well-defined are the Program's measures of outcomes for the STS Research Program?

Factors to consider: The extent to which the Program's products met stakeholder needs in a timely and useful way and ultimately lead to improvements in human and environmental health. The extent to which the program performance measures and LTGs and priorities of the STS MYP link to desired outcomes and Agency priorities. What is the likelihood that the application of products and knowledge by clients would lead toward the achievement of Program outcomes?

(B) Summary Assessment (rate program performance by LTG): A summary assessment and narrative should be provided for LTGs 2 and 3 in the STS MYP. The assessment should be based on three of the questions included above and repeated below, with respect to those P2NT Research Program elements that will be continued within the STS Research Program under LTGs 2 and 3. Since no metrics development work was conducted in the P2NT Research Program, no summary assessment is requested for LTG 1. The summary assessment charge questions are:

- How appropriate is the science used to achieve each LTG in the STS MYP? (i.e., Is the Program asking the right questions, or has it been eclipsed by advancements in the field?)
- How good is the scientific quality of the P2NT research products?
- How much have the results from P2NT research products been used by environmental decision-makers to inform decisions and achieve results?

Elements to include for LTG 2:

The appropriateness, quality, and use of P2NT decision support tools and methodologies to inform stakeholder decisions and achieve results. The extent to which ORD is asking the right questions and conducting the right science to provide tools and methodologies that are responsive to the needs of decision-makers.

Elements to include for LTG 3:

The appropriateness, quality, and use of P2NT technologies to inform stakeholder solutions to environmental problems and achieve results. The extent to which ORD is asking the right questions and conducting the right science to provide technologies that are responsive to the needs of decision-makers.

The BOSC Science and Technology for Sustainability Subcommittee will assign a qualitative score that reflects the quality and significance of the research as well as the extent to which the Program is meeting or making measurable progress toward the goal—relative to the evidence provided to the BOSC. The scores should be in the form of the following adjectives that are defined below and intended to promote consistency among BOSC program reviews. The adjectives should be used as part of a narrative summary of the review, so that the context of the rating and the rationale for selecting a particular rating will be transparent. The rating may reflect considerations beyond the summary assessment questions, and will be explained in the narrative. The adjectives to describe progress are:

- Exceptional indicates that the program is meeting all and exceeding some of its goals, both in the quality of the science being produced and the speed at which research result tools and methods are being produced. An exceptional rating also indicates that the program is addressing the right questions to achieve its goals. The review should be specific as to which aspects of the program's performance have been exceptional.

- Exceeds Expectations indicates that the program is meeting all of its goals. It addresses the appropriate scientific questions to meet its goals and the science is competent or better. It exceeds expectations for either the high quality of the science or for the speed at which work products are being produced and milestones met.

- Meets Expectations indicates that the program is meeting most of its goals. Satisfactory programs live up to expectations in terms of addressing the appropriate scientific questions to meet their goals, and that work products are being produced and milestones

are being reached in a timely manner. The quality of the science being done is competent or better.

? Not Satisfactory indicates that the program is failing to meet a substantial fraction of its goals, or if meeting them, that the achievement of milestones is significantly delayed, or that the questions being addressed are inappropriate or insufficient to meet the intended purpose. Questionable science is also a reason for rating a program as unsatisfactory for a particular long-term goal. The review should be specific as to which aspects of a program's performance have been inadequate.

References

1 Budget Data Request 04-31. Executive Office of the President, Office of Management and Budget. March 22, 2004. ACompleting the Program Assessment Rating Tool (PART) for the FY06 Review Process," pp. 50-56.

2 Memorandum for the Heads of Executive Departments and Agencies. Executive Office of the President, Office of Management and Budget. June 5, 2003. AFY 2005 Interagency Research and Development Priorities," pp. 5-10.

3 Evaluating Federal Research under the Government Performance and Results Act (National Research Council, 1999).

4 The House Science Subcommittee. Letter to Dr. Bruce Alberts, President of the National Academy of Sciences, from F. James Sensenbrenner, Jr. and George E. Brown. October 23, 1997.

5 The Government Performance and Results Act: 1997 Government-Wide Implementation Will Be Uneven. U.S. General Accounting Office. (GAO/GGD, 1997).

6 Building a Foundation for Sound Environmental Decisions. (National Research Council, 1997).

7 ARenewing the Compact between Science and Government,@ Stokes, D.E., in 1995 Forum Proceedings, Vannevar Bush IICScience for the 21st Century. Pp. 15-32. Sigma Xi, 1995.

8 Risk Assessment in the Federal Government: Managing the Process. (National Research Council, 1983).

9 Strengthening Science at the U.S. Environmental Protection Agency. (National Research Council, 2000, p. 141).

Appendix 1: Office of Science and Technology Policy (OSTP)/Office of Management and Budget (OMB) Research and Development Investment Criteria

The Relevance, Quality, and Performance criteria apply to all R&D programs. Industry-relevant applied R&D must meet additional criteria. Together, these criteria can be used to assess the need, relevance, appropriateness, quality, and performance of federal R&D programs.

I. Relevance

R&D investments must have clear plans, must be relevant to national priorities, agency missions, relevant fields, and "customer" needs, and must justify their claim on taxpayer resources. Review committees should assess program objectives and goals on their relevance to national needs, "customer" needs, agency missions, and the field(s) of study the program strives to address. For example, the Joint DOE/NSF Nuclear Sciences Advisory Committee's Long Range Plan and the Astronomy Decadal Surveys are the products of good planning processes because they articulate goals and priorities for research opportunities within and across their respective fields. Programs that directly address Presidential priorities may receive special consideration for support, with adequate documentation of their relevance to those priorities.

OMB will work with some programs to identify quantitative metrics to estimate and compare potential benefits across programs with similar goals. Such comparisons may be within an agency or among agencies.

A. **Programs must have complete plans, with clear goals and priorities.** Programs must provide complete plans, which include explicit statements of: specific issues motivating the program; broad goals and more specific tasks meant to address the issues; priorities among goals and activities within the program; human and capital resources anticipated; and intended program outcomes, against which success may later be assessed.

B. **Programs must articulate the potential public benefits of the program.** Programs must identify potential benefits, including added benefits beyond those of any similar efforts that have been or are being funded by the government or others. R&D benefits may include technologies and methods that could provide new options in the future, if the landscape of today's needs and capabilities changes dramatically. Some programs and sub-program units may be required to quantitatively estimate expected benefits, which would include metrics to permit meaningful comparisons among programs that promise similar benefits. While all programs should try to articulate potential benefits, OMB and OSTP recognize the difficulty in predicting the outcomes of basic research. Discovery is a legitimate object of basic research, and some basic research investments may be justified on external judgments of the opportunity for discovery.

C. **Programs must document their relevance to specific Presidential priorities to receive special consideration.** Many areas of research warrant some level of federal funding. Nonetheless, the President has identified a few specific areas of research that are particularly important. To the extent that a proposed project can document how it directly addresses one of these areas, it may be given preferential treatment.

D. Program relevance to the needs of the Nation, of fields of science and technology, and of program "customers" must be assessed through prospective external review. Programs must be assessed on their relevance to agency missions, fields of science or technology, or other "customer" needs. A customer may be another program at the same or another agency, an interagency initiative or partnership, or a firm or other organization from another sector or country. As appropriate, programs must define a plan for regular reviews by primary customers of the program's relevance to their needs. These programs must provide a plan for addressing the conclusions of external reviews.

E. Program relevance to the needs of the Nation, of fields of science and technology, and of program "customers" must be assessed periodically through retrospective external review. Programs must periodically assess the need for the program and its relevance to customers against the original justifications. Programs must provide a plan for addressing the conclusions of external reviews.

II. Quality

Programs should maximize the quality of the R&D they fund through the use of a clearly stated, defensible method for awarding a significant majority of their funding. A customary method for promoting R&D quality is the use of a competitive, merit-based process. The National Science Foundation's (NSF) process for the peer-reviewed, competitive award of its R&D grants is a good example. Justifications for processes other than competitive merit review may include "outside-the-box" thinking, a need for timeliness (e.g., R&D grants for rapid studies in response to an emergency), unique skills or facilities, or a proven record of outstanding performance (e.g., performance-based renewals).

Programs must assess and report on the quality of current and past R&D. For example, NSF's use of Committees of Visitors, which review NSF directorates, is an example of a good quality-assessment tool. OMB and OSTP encourage agencies to provide the means by which their programs may be benchmarked internationally or across agencies, which provides one indicator of program quality.

A. Programs allocating funds through means other than a competitive, merit-based process must justify funding methods and document how quality is maintained. Programs must clearly describe how much of the requested funding will be broadly competitive based on merit, providing compelling justifications for R&D funding allocated through other means. (See OMB Circular A-11 for definitions of competitive merit review and other means of allocating federal research funding.) All program funds allocated through means other than unlimited competition must document the processes they will use to distribute funds to each type of R&D performer (e.g., federal laboratories, federally funded R&D centers, universities). Programs are encouraged to use external assessment of the methods they use to allocate R&D and maintain program quality.

B. Program quality must be assessed periodically through retrospective expert review. Programs must institute a plan for regular, external reviews of the quality of the program's research and research performers, including a plan to use the results from these reviews to guide future program decisions. Rolling reviews performed every 3-5 years by

advisory committees can satisfy this requirement. Benchmarking of scientific leadership and other factors provides an effective means of assessing program quality relative to other programs, other agencies, and other countries.

III. Performance

R&D programs should maintain a set of high priority, multi-year R&D objectives with annual performance measures and milestones that show how one or more outcomes will be reached. Metrics should be defined not only to encourage individual program performance but also to promote, as appropriate, broader goals, such as innovation, cooperation, education, and dissemination of knowledge, applications, or tools.

OMB encourages agencies to make the processes they use to satisfy the Government Performance and Results Act (GPRA) consistent with the goals and metrics they use to satisfy these R&D criteria. Satisfying the R&D performance criteria for a given program should serve to set and evaluate R&D performance goals for the purposes of GPRA. OMB expects goals and performance measures that satisfy the R&D criteria to be reflected in agency performance plans.

Programs must demonstrate an ability to manage in a manner that produces identifiable results. At the same time, taking risks and working towards difficult-to-attain goals are important aspects of good research management, especially for basic research. The intent of the investment criteria is not to drive basic research programs to pursue less risky research that has a greater chance of success. Instead, the Administration will focus on improving the management of basic research programs.

OMB will work with some programs to identify quantitative metrics to compare performance across programs with similar goals. Such comparisons may be within an agency or among agencies.

Construction projects and facility operations will require additional performance metrics. Cost and schedule earned-value metrics for the construction of R&D facilities must be tracked and reported. Within DOE, the Office of Science's formalized independent reviews of technical cost, scope, and schedule baselines and project management of construction projects ("Lehman Reviews") are widely recognized as an effective practice for discovering and correcting problems involved with complex, one-of-a-kind construction projects.

A. Programs may be required to track and report relevant program inputs annually. Programs may be expected to report relevant program inputs, which could include statistics on overhead, intramural/extramural spending, infrastructure, and human capital. These inputs should be discussed with OMB.

B. Programs must define appropriate output and outcome measures, schedules, and decision points. Programs must provide single- and multi-year R&D objectives, with annual performance measures, to track how the program will improve scientific understanding and its application. Programs must provide schedules with annual milestones for future competitions, decisions, and termination points, highlighting

changes from previous schedules. Program proposals must define what would be a minimally effective program and a successful program. Agencies should define appropriate output and outcome measures for all R&D programs, but agencies should not expect fundamental basic research to be able to identify outcomes and measure performance in the same way that applied research or development are able to. Highlighting the results of basic research is important, but it should not come at the expense of risk-taking and innovation. For some basic research programs, OMB may accept the use of qualitative outcome measures and quantitative process metrics. Facilities programs must define metrics and methods (e.g., earned-value reporting) to track development costs and to assess the use and needs of operational facilities over time. If leadership in a particular field is a goal for a program or agency, OMB and OSTP encourage the use of benchmarks to assess the processes and outcomes of the program with respect to leadership. OMB encourages agencies to make the processes they use to satisfy GPRA consistent with the goals and metrics they use to satisfy these R&D criteria.

C. **Program performance must be retrospectively documented annually.** Programs must document performance against previously defined output and outcome metrics, including progress towards objectives, decisions, and termination points or other transitions. Programs with similar goals may be compared on the basis of their performance. OMB will work with agencies to identify such programs and appropriate metrics to enable such comparisons.

IV. Criteria for R&D Programs Developing Technologies That Address Industry Issues

The purpose of some R&D and technology demonstration programs and projects is to introduce some product or concept into the marketplace. However, some of these efforts engage in activities that industry is capable of doing and may discourage or even displace industry investment that would otherwise occur. Programs should avoid duplicating research in areas that are receiving funding from the private sector, especially for evolutionary advances and incremental improvements. For the purposes of assessing federal R&D investments, the following criteria should be used to assess industry-relevant R&D and demonstration projects, including, at OMB discretion, associated construction activities.

OMB will work with programs to identify appropriate measures to compare potential benefits and performance across programs with similar goals, as well as ways to assess market relevance.

A. **Programs and projects must articulate public benefits of the program using uniform benefit indicators across programs and projects with similar goals.** In addition to the public benefits required in the general criteria, all industry-relevant programs and projects must identify and use uniform benefit indicators (including benefit-cost ratios) to enable comparisons of expected benefits across programs and projects. OMB will work with agencies to identify these indicators.

B. **Programs and projects must justify the appropriateness of federal investment.** Programs and projects must demonstrate that industry investment is sub-optimal to

develop a technology or system and explain why the development or acceleration of that technology or system is necessary to meet a federal mission or goals.

C. **Programs and projects must demonstrate that investment in R&D and demonstration activities is a more effective way to support the federal goals than other policy alternatives.** When the federal government chooses to intervene to address market failures, there may be many policy alternatives to address those failures. Among other tools available to the government are legislation, tax policy, regulatory and enforcement efforts, and an integrated combination of these approaches. Agencies should consider that the legislation, tax policy or regulatory or enforcement mechanisms may already be in place to achieve a reasonable expectation of advancing the desired end.

D. **Programs and projects must document industry or market relevance, including readiness of the market to adopt technologies or other outputs.** Programs must assess the likelihood that the target industry will be able to adopt the technology or other program outputs. The level of industry cost sharing or enforceable recoupment commitments in contracts are indicators of industry relevance. Agencies must be able to justify any demonstration activities with an economic analysis of the public and private returns on the public investment.

E. **Program performance plans and reports must include "off ramps" and transition points.** In addition to the schedules and decision points defined in the general criteria, program plans should also identify whether, when, and how aspects of the program may be shifted to the private sector.

Appendix D: List of Acronyms

APG	Annual Performance Goal
BEES	Building for Environmental and Economic Sustainability
BOSC	Board of Scientific Counselors
CNS	Collaborative Science and Technology Network for Sustainability
DOE	U.S. Department of Energy
EPA	U.S. Environmental Protection Agency
ETV	Environmental Technology Verification
FACA	Federal Advisory Committee Act
GAO	General Accountability Office
GPRA	Government Performance and Results Act
HPDP	Hierarchical Patch Dynamics Paradigm
ISO	International Organization for Standardization
LCA	Life Cycle Assessment
LTG	Long-Term Goal
MYP	Multi-Year Plan
NGO	Nongovernmental Organization
NIST	National Institute of Standards and Technology
NSF	National Science Foundation
OMB	Office of Management and Budget
ORD	Office of Research and Development
OSTP	Office of Science and Technology Policy
P2NT	Pollution Prevention and New Technologies
P3	People, Prosperity, and the Planet
PART	Program Assessment Rating Tool
SAB	Science Advisory Board
SBIR	Small Business Innovation Research
SETAC	Society of Environmental Toxicology and Chemistry
SMART	Specific, Measurable, Achievable, Relevant, Timely
STS	Science and Technology for Sustainability
TRACI	Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts
TSE	Technology for a Sustainable Environment
USPTO	U.S. Patent and Trademark Office